

#5

OIPE

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/617,389  
DATE: 07/28/2000  
TIME: 14:36:07Input Set : A:\9606C3.seq.txt  
Output Set: N:\CRF3\07282000\I617389.raw

ENTERED

```
3 <110> APPLICANT: Lok, Si
4 Conklin, Darrell C.
5 Adams, Robyn L.
6 Jeimberg, Anna C.
7 Lofton-Day, Catherine E.
8 Jaspers, Stephen R.
9 Stamm, Mike R.
11 <120> TITLE OF INVENTION: TESTIS-SPECIFIC INSULIN HOMOLOG POLYPEPTIDES
13 <130> FILE REFERENCE: 96-06C3
C--> 15 <140> CURRENT APPLICATION NUMBER: US/09/617,389
C--> 15 <141> CURRENT FILING DATE: 2000-07-17
15 <150> PRIOR APPLICATION NUMBER: 08/905,267
16 <151> PRIOR FILING DATE: 1997-08-01
18 <150> PRIOR APPLICATION NUMBER: 60/023,213
19 <151> PRIOR FILING DATE: 1996-08-02
21 <150> PRIOR APPLICATION NUMBER: 60/031,592
22 <151> PRIOR FILING DATE: 1996-11-21
24 <160> NUMBER OF SEQ ID NOS: 15
26 <170> SOFTWARE: FastSEQ for Windows Version 3.0
28 <210> SEQ ID NO: 1
29 <211> LENGTH: 566
30 <212> TYPE: DNA
31 <213> ORGANISM: Rattus norvegicus
33 <220> FEATURE:
34 <221> NAME/KEY: CDS
35 <222> LOCATION: (1)...(566)
37 <400> SEQUENCE: 1
38 atg aag cag ctg tgc tgt tct tgt ctg ttg tgg ctt gga ctc cta ctg 48
39 Met Lys Gln Leu Cys Cys Ser Cys Leu Leu Trp Leu Gly Leu Leu Leu
40 1 5 10 15
42 gct cct ttc tcc cag gaa caa gaa gag gtc acc agc ccc acg aag ttg 96
43 Ala Pro Phe Ser Gln Glu Gln Glu Glu Val Thr Ser Pro Thr Lys Leu
44 20 25 30
46 tgc ggc agg gac ctg ttg gta gaa gtt ata aaa ctc tgt ggc caa aat 144
47 Cys Gly Arg Asp Leu Leu Val Glu Val Ile Lys Leu Cys Gly Gln Asn
48 35 40 45
50 gac tgg agc cgg ttc tcg atg gaa gag caa agt cct atg aca gag ttg 192
51 Asp Trp Ser Arg Phe Ser Met Glu Glu Gln Ser Pro Met Thr Glu Leu
52 50 55 60
54 gtt ccc caa tat aca cgg aaa gtc aaa acc ttc aac cct cac cgg tcc 240
55 Val Pro Gln Tyr Thr Arg Lys Val Lys Thr Phe Asn Pro His Arg Ser
56 65 70 75 80
58 tcc tcc tcc tgg gga aga ttc aca aac cca ggc gtc tcc cag aag aaa 288
59 Ser Ser Ser Trp Gly Arg Phe Thr Asn Pro Gly Val Ser Gln Lys Lys
60 85 90 95
62 gca aca cac act tgg gaa tct cag tca ctg ccc aac tat cag ctt aaa 336
63 Ala Thr His Thr Trp Glu Ser Gln Ser Leu Pro Asn Tyr Gln Leu Lys
```

RAW SEQUENCE LISTING                      DATE: 07/28/2000  
 PATENT APPLICATION:    US/09/617,389                      TIME: 14:36:07

Input Set : A:\9606C3.seq.txt  
 Output Set: N:\CRF3\07282000\I617389.raw

```

64          100          105          110
66 aag gag gag ctg ctt ccg aag aca gga gtg cat tca tac cac ggt ggc      384
67 Lys Glu Glu Leu Leu Pro Lys Thr Gly Val His Ser Tyr His Gly Gly
68          115          120          125
70 aag ccc tat gtg aag agt gta aaa ttt cag aag aaa aac act gac aaa      432
71 Lys Pro Tyr Val Lys Ser Val Lys Phe Gln Lys Lys Asn Thr Asp Lys
72          130          135          140
74 atg agt acc ttc agc ggc tta ttt tgg ggg aac cat ccc cag agg aag      480
75 Met Ser Thr Phe Ser Gly Leu Phe Trp Gly Asn His Pro Gln Arg Lys
76          145          150          155          160
78 cgc aga ggt ttc gca gat aaa tgc tgt gct ata ggg tgc tcc aaa gag      528
79 Arg Arg Gly Phe Ala Asp Lys Cys Cys Ala Ile Gly Cys Ser Lys Glu
80          165          170          175
82 gag ctg gcc gtc gca tgc ctt ccg ttt gtt gat ttt ta      566
W--> 83 Glu Leu Ala Val Ala Cys Leu Pro Phe Val Asp Phe
84          180          185
87 <210> SEQ ID NO: 2
88 <211> LENGTH: 188
89 <212> TYPE: PRT
90 <213> ORGANISM: Rattus norvegicus
92 <400> SEQUENCE: 2
93 Met Lys Gln Leu Cys Cys Ser Cys Leu Leu Trp Leu Gly Leu Leu Leu
94   1          5          10          15
95 Ala Pro Phe Ser Gln Glu Gln Glu Val Thr Ser Pro Thr Lys Leu
96          20          25          30
97 Cys Gly Arg Asp Leu Leu Val Glu Val Ile Lys Leu Cys Gly Gln Asn
98          35          40          45
99 Asp Trp Ser Arg Phe Ser Met Glu Glu Gln Ser Pro Met Thr Glu Leu
100          50          55          60
101 Val Pro Gln Tyr Thr Arg Lys Val Lys Thr Phe Asn Pro His Arg Ser
102          65          70          75          80
103 Ser Ser Ser Trp Gly Arg Phe Thr Asn Pro Gly Val Ser Gln Lys Lys
104          85          90          95
105 Ala Thr His Thr Trp Glu Ser Gln Ser Leu Pro Asn Tyr Gln Leu Lys
106          100          105          110
107 Lys Glu Glu Leu Leu Pro Lys Thr Gly Val His Ser Tyr His Gly Gly
108          115          120          125
109 Lys Pro Tyr Val Lys Ser Val Lys Phe Gln Lys Lys Asn Thr Asp Lys
110          130          135          140
111 Met Ser Thr Phe Ser Gly Leu Phe Trp Gly Asn His Pro Gln Arg Lys
112          145          150          155          160
113 Arg Arg Gly Phe Ala Asp Lys Cys Cys Ala Ile Gly Cys Ser Lys Glu
114          165          170          175
115 Glu Leu Ala Val Ala Cys Leu Pro Phe Val Asp Phe
116          180          185
118 <210> SEQ ID NO: 3
119 <211> LENGTH: 49
120 <212> TYPE: DNA
121 <213> ORGANISM: Artificial Sequence

```

RAW SEQUENCE LISTING                      DATE: 07/28/2000  
 PATENT APPLICATION: US/09/617,389        TIME: 14:36:07

Input Set : A:\9606C3.seq.txt  
 Output Set: N:\CRF3\07282000\I617389.raw

```

123 <220> FEATURE:
124 <223> OTHER INFORMATION: Oligonucleotide zc6091
126 <400> SEQUENCE: 3
127 gagcacagaa ttcactactc gaggcggccg cttttttttt tttttttt      49
129 <210> SEQ ID NO: 4
130 <211> LENGTH: 32
131 <212> TYPE: DNA
132 <213> ORGANISM: Artificial Sequence
134 <220> FEATURE:
135 <223> OTHER INFORMATION: Oligonucleotide ZC10008
137 <400> SEQUENCE: 4
138 caggtcgaca ccatgccgcg gctcctccgc tt                        32
140 <210> SEQ ID NO: 5
141 <211> LENGTH: 34
142 <212> TYPE: DNA
143 <213> ORGANISM: Artificial Sequence
145 <220> FEATURE:
146 <223> OTHER INFORMATION: Oligonucleotide ZC10491
148 <400> SEQUENCE: 5
149 gagtctagat tatgtcactc atacaggaaa gctc                        34
151 <210> SEQ ID NO: 6
152 <211> LENGTH: 23
153 <212> TYPE: DNA
154 <213> ORGANISM: Artificial Sequence
156 <220> FEATURE:
157 <223> OTHER INFORMATION: Oligonucleotide ZC10684
159 <400> SEQUENCE: 6
160 tcctccgctt gtcctgctg tgg                                23
162 <210> SEQ ID NO: 7
163 <211> LENGTH: 23
164 <212> TYPE: DNA
165 <213> ORGANISM: Artificial Sequence
167 <220> FEATURE:
168 <223> OTHER INFORMATION: Oligonucleotide ZC10685
170 <400> SEQUENCE: 7
171 ctctctgaaa cggaactggc tcc                                23
173 <210> SEQ ID NO: 8
174 <211> LENGTH: 574
175 <212> TYPE: DNA
176 <213> ORGANISM: Artificial Sequence
178 <220> FEATURE:
179 <223> OTHER INFORMATION: Human testis EST
181 <400> SEQUENCE: 8
182 gtcgacacca tgccgcggct cctccgcttg tccctgctgt ggcttggact cctgctggtt      60
183 cggttttctc gtgaactgag cgacatcagc agtgccagga agctgtgcgg cagggtacttg      120
184 gtgaaaagaaa tagaaaaact ctgcggccat gccaaactgga gccagtccg ttctgaggag      180
185 gaaacccctt tctcacggtt gattgcacag gcctcggaga aggtcgaagc ctacagccca      240
186 taccagttcg aaagcccgcg aaccgcttcc ccggcccggg gaagaggcac aaacccagtg      300
187 tctactttctt gggaagaagc agtaaacagt tgggaaatgc agtcactacc tgagtataag      360

```

RAW SEQUENCE LISTING                      DATE: 07/28/2000  
 PATENT APPLICATION: US/09/617,389              TIME: 14:36:07

Input Set : A:\9606C3.seq.txt  
 Output Set: N:\CRF3\07282000\I617389.raw

```

188 gataaaaagg gatattcacc ccttggtctg ctggagtttg ctggaggccc actccagatg      420
189 ctgtttgcct gggatcacc agcagaggct gcagaacagc aaagattgct gcctgttctc      480
190 tcctctggaa gcttcacccc agagggggcac ccactagatg ccagccagag ctttctctga      540
191 tgagtgacat aaggattcaa cttcaacaga attc                                574
193 <210> SEQ ID NO: 9
194 <211> LENGTH: 26
195 <212> TYPE: DNA
196 <213> ORGANISM: Artificial Sequence
198 <220> FEATURE:
199 <223> OTHER INFORMATION: Oligonucleotide ZC10491
201 <400> SEQUENCE: 9
202 cgaccgccat tgcacaacgc ggagga                                26
204 <210> SEQ ID NO: 10
205 <211> LENGTH: 35
206 <212> TYPE: DNA
207 <213> ORGANISM: Artificial Sequence
209 <220> FEATURE:
210 <223> OTHER INFORMATION: Oligonucleotide ZC10537
212 <400> SEQUENCE: 10
213 ataggaggaa aggttggttg tgacaagaa acttg                                35
215 <210> SEQ ID NO: 11
216 <211> LENGTH: 34
217 <212> TYPE: DNA
218 <213> ORGANISM: Artificial Sequence
220 <220> FEATURE:
221 <223> OTHER INFORMATION: Oligonucleotide ZC10538
223 <400> SEQUENCE: 11
224 gacgaattct gttgaagttg aatccttatg tcac                                34
226 <210> SEQ ID NO: 12
227 <211> LENGTH: 703
228 <212> TYPE: DNA
229 <213> ORGANISM: Homo sapiens
231 <220> FEATURE:
232 <221> NAME/KEY: CDS
233 <222> LOCATION: (17)...(658)
235 <400> SEQUENCE: 12
236 gcctggggtc acaggg atg ccg cgg ctc ctc cgc ttg tcc ctg ctg tgg ctt      52
237 Met Pro Arg Leu Leu Arg Leu Ser Leu Leu Trp Leu
238      1      5      10
240 gga ctc ctg ctg gtt cgg ttt tct cgt gaa ctg agc gac atc agc agt      100
241 Gly Leu Leu Leu Val Arg Phe Ser Arg Glu Leu Ser Asp Ile Ser Ser
242      15      20      25
244 gcc agg aag ctg tgc ggc agg tac ttg gtg aaa gaa ata gaa aaa ctc      148
245 Ala Arg Lys Leu Cys Gly Arg Tyr Leu Val Lys Glu Ile Glu Lys Leu
246      30      35      40
248 tgc ggc cat gcc aac tgg agc cag ttc cgt ttc gag gag gaa acc cct      196
249 Cys Gly His Ala Asn Trp Ser Gln Phe Arg Phe Glu Glu Glu Thr Pro
250      45      50      55      60
252 ttc tca cgg ttg att gca cag gcc tcg gag aag gtc gaa gcc tac agc      244

```

RAW SEQUENCE LISTING      DATE: 07/28/2000  
 PATENT APPLICATION: US/09/617,389      TIME: 14:36:07

Input Set : A:\9606C3.seq.txt  
 Output Set: N:\CRF3\07282000\I617389.raw

```

253 Phe Ser Arg Leu Ile Ala Gln Ala Ser Glu Lys Val Glu Ala Tyr Ser
254          65          70          75
256 cca tac cag ttc gaa agc ccg caa acc gct tcc ccg gcc cgg gga aga      292
257 Pro Tyr Gln Phe Glu Ser Pro Gln Thr Ala Ser Pro Ala Arg Gly Arg
258          80          85          90
260 ggc aca aac cca gtg tct act tct tgg gaa gaa gca gta aac agt tgg      340
261 Gly Thr Asn Pro Val Ser Thr Ser Trp Glu Glu Ala Val Asn Ser Trp
262          95          100          105
264 gaa atg cag tca cta cct gag tat aag gat aaa aag gga tat tca ccc      388
265 Glu Met Gln Ser Leu Pro Glu Tyr Lys Asp Lys Lys Gly Tyr Ser Pro
266          110          115          120
268 ctt ggt aag aca aga gaa ttt tct tca tca cat aat atc aat gta tat      436
269 Leu Gly Lys Thr Arg Glu Phe Ser Ser Ser His Asn Ile Asn Val Tyr
270          125          130          135          140
272 att cat gag aat gca aaa ttt cag aag aaa cgt aga aac aaa att aaa      484
273 Ile His Glu Asn Ala Lys Phe Gln Lys Lys Arg Arg Asn Lys Ile Lys
274          145          150          155
276 acc tta agc aat ttg ttt tgg ggg cat cat ccc caa aga aaa cgc aga      532
277 Thr Leu Ser Asn Leu Phe Trp Gly His His Pro Gln Arg Lys Arg Arg
278          160          165          170
280 gga tat tca gaa aag tgt tgt ctt aca gga tgt aca aaa gaa gaa ctt      580
281 Gly Tyr Ser Glu Lys Cys Cys Leu Thr Gly Cys Thr Lys Glu Glu Leu
282          175          180          185
284 agc att gca tgt ctt cca tat att gat ttt aaa agg cta aag gaa aaa      628
285 Ser Ile Ala Cys Leu Pro Tyr Ile Asp Phe Lys Arg Leu Lys Glu Lys
286          190          195          200
288 aga tca tca ctt gta act aag ata tac taa ccattcttaga attttttcta      678
289 Arg Ser Ser Leu Val Thr Lys Ile Tyr *
290          205          210
292 acctaataaa agcttaatac attta      703
294 <210> SEQ ID NO: 13
295 <211> LENGTH: 213
296 <212> TYPE: PRT
297 <213> ORGANISM: Homo sapiens
299 <400> SEQUENCE: 13
300 Met Pro Arg Leu Leu Arg Leu Ser Leu Leu Trp Leu Gly Leu Leu Leu
301          1          5          10          15
302 Val Arg Phe Ser Arg Glu Leu Ser Asp Ile Ser Ser Ala Arg Lys Leu
303          20          25          30
304 Cys Gly Arg Tyr Leu Val Lys Glu Ile Glu Lys Leu Cys Gly His Ala
305          35          40          45
306 Asn Trp Ser Gln Phe Arg Phe Glu Glu Glu Thr Pro Phe Ser Arg Leu
307          50          55          60
308 Ile Ala Gln Ala Ser Glu Lys Val Glu Ala Tyr Ser Pro Tyr Gln Phe
309          65          70          75          80
310 Glu Ser Pro Gln Thr Ala Ser Pro Ala Arg Gly Arg Gly Thr Asn Pro
311          85          90          95
312 Val Ser Thr Ser Trp Glu Glu Ala Val Asn Ser Trp Glu Met Gln Ser
313          100          105          110

```

VERIFICATION SUMMARY                      DATE: 07/28/2000  
PATENT APPLICATION:    US/09/617,389        TIME: 14:36:08

Input Set : A:\9606C3.seq.txt  
Output Set: N:\CRF3\07282000\I617389.raw

L:15 M:270 C: Current Application Number differs, Replaced Current Application No  
L:15 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:83 M:361 W: Invalid Split Codon, Sequence data for SEQ ID#: 1